

11/2  
03-d-d

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/508,821A

DATE: 02/23/2001  
TIME: 16:03:53

Input Set : A:\Pto.amc  
Output Set: N:\CRF3\02232001\I508821A.raw

2 <110> APPLICANT: ROULEAU, Guy A.  
 3 JOOBER, Ridha  
 4 BENKELFAT, Chawki  
 6 <120> TITLE OF INVENTION: POLYMORPHIC CAG REPEAT-CONTAINING GENE AND USES THEREOF  
 8 <130> FILE REFERENCE: 12948-10  
 10 <140> CURRENT APPLICATION NUMBER: 09/508,821A  
 11 <141> CURRENT FILING DATE: 2000-05-26  
 13 <150> PRIOR APPLICATION NUMBER: PCT/CA98/00884  
 14 <151> PRIOR FILING DATE: 1998-09-18  
 16 <150> PRIOR APPLICATION NUMBER: 2,216,057 CA  
 17 <151> PRIOR FILING DATE: 1997-09-19  
 19 <160> NUMBER OF SEQ ID NOS: 6  
 21 <170> SOFTWARE: PatentIn Ver. 2.1  
 23 <210> SEQ ID NO: 1  
 24 <211> LENGTH: 328  
 25 <212> TYPE: DNA  
 26 <213> ORGANISM: Mus musculus  
 28 <400> SEQUENCE: 1  
 29 agggcagcca ctttccccag cattcccaact ctctccatc tatgccccaa 60  
 30 cagtgcaggg tgggtggcagq ggggccccact octacaagag ctgcacagca ccatactgccc 120  
 31 agcctcatga taggcgcattg agtgccatcg cgaacctggc ttccaggcaga cgggtccaga 180  
 32 atcttcacgc ttaccagctt ggcgccttg gctacagagca gcacgcgcata qcacttcaag 240  
 33 gccgtcacca cacccagggaa acactccact accagaacct cgccaaatgc caaacatcg 300  
 34 gacagcaagg ccagggtctac tgccacc 328  
 37 <210> SEQ ID NO: 2  
 38 <211> LENGTH: 330  
 39 <212> TYPE: DNA  
 40 <213> ORGANISM: Homo sapiens  
 42 <220> FEATURE:  
 43 <221> NAME/KEY: modified\_base  
 44 <222> LOCATION: (59)  
 45 <223> OTHER INFORMATION: N = A or G or C or T/U  
 47 <400> SEQUENCE: 2  
 W--> 48 tccttccca cttccatcc actactctcc tctgtcccaag ggtgggtggc agggggccna 60  
 49 ctccataaag aatgcacag caccgactgc ccagcccatc gacaggccgc tgactgccag 120  
 50 ctccatcgcc gccccggggc aeggggtcca gaatcttcat gcttaccatg cggggccct 180  
 51 cagctatgac cagcagcagc aqccacgagca gcacgcgcagc cagcagcgc aageccctca 240  
 52 gagccggcac catccccagg aaacccttca ttacccaaaac ctgcgcattt atcagcacta 300  
 53 cgggcagcaa ggccagggtct actgcgcagcc 330  
 56 <210> SEQ ID NO: 3  
 57 <211> LENGTH: 23  
 58 <212> TYPE: DNA  
 59 <213> ORGANISM: Artificial Sequence  
 61 <220> FEATURE:  
 62 <223> OTHER INFORMATION: Description of Artificial Sequence:  
 63 oligonucleotide  
 65 <400> SEQUENCE: 3

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/508,821A

DATE: 02/23/2001  
TIME: 16:03:53

Input Set : A:\Pto.amc  
Output Set: N:\CRF3\02232001\1508821A.raw

66 ggccggcagg gtccagaata ttc 23  
 69 <210> SEQ ID NO: 4  
 70 <211> LENGTH: 23  
 71 <212> TYPE: DNA  
 72 <213> ORGANISM: Artificial Sequence  
 74 <220> FEATURE:  
 75 <223> OTHER INFORMATION: Description of Artificial Sequence:  
 76 oligonucleotide  
 78 <400> SEQUENCE: 4 23  
 79 agcactacgg gcagcaaggc cag  
 82 <210> SEQ ID NO: 5  
 83 <211> LENGTH: 6022  
 84 <212> TYPE: DNA  
 85 <213> ORGANISM: Homo sapiens  
 87 <400> SEQUENCE: 5  
 88 ggatccaggca gccccaaagg gatgaggggc cgaaaaattgc tctgtctaaat gcttttggac 60  
 89 tglcaggaaq ggctgggggt gatgggtggg ggacattggg gaggagctgg caatggcgq 120  
 90 ggggggggggg gglagctccc ctagtgcacggc ggcgtggcga gccgggttttgc ctcggcgat 180  
 91 cagtggccgt ccttggcaag acteagctgc aggccatgtg ggagcggggaa ttacagagca 240  
 92 caccctccgt acacagaatg tglcaataatq cgcacacgtg gtggggggggc tcaggcgaag 300  
 93 ggggggactat taaqagctgc gggggggggc aggccagggtggg gggagggtggg tggggagggtg 360  
 94 ctttctgagg caaaaggaaatggcccgatc gaatecgctca ttcctctggcc ctcctcgcc 420  
 95 ctcctccccc tcctctcttc ctcctccccc tccttcctt tttcttttca cagataacc 480  
 96 gccccggatca tgcagttttt tccgayaagg tgggggttcc atggccaaaca acagaactac 540  
 97 cccggggactt cccggggaaat atccgccta gagaatttaca ggcggccggag tccggccgg 600  
 98 cttaatgtggc accggggcggc gctgtctggc aaggactattt ataacccgcg gctttaacc 660  
 99 agctatgggg gtggcccgatc cccggccctt ggcactgeag ccgggggtggc cggccgacaa 720  
 100 taccacccgg gcaagcaaggc cctggcccaaca cagcaaggcc tgcaggggggq gccggctt 780  
 101 cctggcttcg ggcgcaggaa cccggccccc taccggggcc gctatgttgg tgaggagac 840  
 102 ctccgggtt gggggggccccccca aacggccatcca ccccccacccggccggccact acctggcagg 900  
 103 gtggccaaat atgtatggaaat cttgtatggaa aagacacggc tgcggcccccggc caggccgtat 960  
 104 gcaagacggg ggcggccggat ggccttccgg actcactccc tgcacgttcc gcaaggccacq 1020  
 105 cccggcccccggc aacggccatggc atcccaaaaggc acaatggcggatc gaaagacatt 1080  
 106 ggcggccccc tgcggccccc cccgggttcc cccggccccc gcatatccca gtccttcggcc 1140  
 107 accttcctca ccttacttcc ctcgtggccggc gggggggccca ctcctataaag 1200  
 108 agttgcacag cccggccatcc cccggccatcc gacaggccggc tgcacgttcc ctcggccct 1260  
 109 gccccggggc aacggggatc gatgttccat qctatccatq cggggccggccct cagctatggc 1320  
 110 cccggcccccggc aacggccatggc gacggccatcc gacggccatcc aacggccatcc gggccggccac 1380  
 111 catggccgggg cccggccatcc ttacccaaac ctcggccaaat atcggccatca cggccggccaa 1440  
 112 ggcggccccc gtcggccatcc gacggccatcc gtcggccatcc ctcggccatcc gtcggccatcc 1500  
 113 ttcggccatcc gtcggccatcc ctcggccatcc gtcggccatcc gtcggccatcc ttcggccatcc 1560  
 114 tccacccatcc gtcggccatcc gtcggccatcc gtcggccatcc gtcggccatcc gtcggccatcc 1620  
 115 aacggcccccggc ctcggccatcc gtcggccatcc gtcggccatcc gtcggccatcc gtcggccatcc 1680  
 116 cccggccatcc gtcggccatcc gtcggccatcc gtcggccatcc gtcggccatcc gtcggccatcc 1740  
 117 cagaaggaca agtcgttcc gtcggccatcc gtcggccatcc gtcggccatcc gtcggccatcc 1800  
 118 ctggccatcc gtcggccatcc gtcggccatcc gtcggccatcc gtcggccatcc gtcggccatcc 1860  
 119 gtcggccatcc gtcggccatcc gtcggccatcc gtcggccatcc gtcggccatcc gtcggccatcc 1920  
 120 aacggccatcc gtcggccatcc gtcggccatcc gtcggccatcc gtcggccatcc gtcggccatcc 1980  
 121 tcggccatcc gtcggccatcc gtcggccatcc gtcggccatcc gtcggccatcc gtcggccatcc 2040

RAW SEQUENCE LISTING DATE: 02/23/2001  
PATENT APPLICATION: US/09/508,821A TIME: 16:03:53

Input Set : A:\Pto.amc  
Output Set: N:\CRF3\02232001\1508821A.raw

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/508,821A

DATE: 02/23/2001  
TIME: 16:03:53

Input Set : A:\Pto.amc  
Output Set: N:\CRF3\02232001\I508821A.raw

```

171 ccaaaggatgg aggagctggg ccctgcclcc cagccccccgg agggcaggcc ctgccagcc 5040
172 cagaccaaggg cacagaaaca gcccggccac accaactaca gcagtttattc caagcgaaag 5100
173 cgcctcaactt gggggccgggg caaqaacacc acctcttccac ctgtataaggg gctgtccaaq 5160
174 cgaacgacacgaa cccggccgggg gtcggccctg gatccggccag aacctgaaaal ccgcctcaaq 5220
175 tacattttctt ctgtcaagcg gctgagggtca gacagccggaa ccccccgtt ctcacccttc 5280
176 gtgcgggtgg agaaggcgaga cgcgttcacc accatgtca ctgttgtcaa ctcccttgaa 5340
177 gatgcggccaa agcccccacag qaagecttcc tcctctgtctt ctcttcttc atctctgtcc 5400
178 tggttctctt tggtatgtcaqc cggggccctcc ctggccacac tccctggagg ctccatctg 5460
179 cagccggccggtt cctcttcttc ctcttcttc atgttcactt tggtggctgt ggtttccaaq 5520
180 qcclgagta ctcttgcgt tggttgtcc tccttgcctaa accccggccaa cttaaggac 5580
181 ctggggacc tctgtggcc ctactaccctt gaacacttcc ttccccaaaaaa qaagccaaaa 5640
182 ctcaaggaga aggtgcggcc agaaggccacc ttgtggggggg ctctgtccgc qcttggagata 5700
183 acactcaaaag gtcggcgggtg tgcactgtcc gcaacttccgg ggaagcccccc cagggtgacgg 5760
184 cccagtcgtac ccggccaaqcc agggcccaactt ggcgcaccatgt gccccggccg tggccggag 5820
185 gtcgcagacg tgcactgtctt gtgtatggccgg ggaggatgggg ggegaggagg cagcccccc 5880
186 cggacaagggtt cccaaacatq agtcgcggccgg gggaggctccq gcaagcccccc qcgqqqgq 5940
187 ccaaggacac tgggtgcatttggggccatg aggctgttcc gtcgttggacc gggggctt accttgggtgc 6000
188 cggggaaatc ttggggcttc ag 6022
191 <210> SEQ ID NO: 6
192 <211> LENGTH: 22
193 <212> TYPE: DNA
194 <213> ORGANISM: Artificial Sequence
195 <220> FEATURE:
196 <223> OTHER INFORMATION: Description of Artificial Sequence:
197 oligonucleotide
200 <400> SEQUENCE: 6
201 tggccttgtcttggccatgttggacc
```

22

VERIFICATION SUMMARY  
PATENT APPLICATION: US/09/508,821A DATE: 02/23/2001  
TIME: 16:03:54

Input Set : A:\pto.amc  
Output Set: N:\CRF3\02232001\I508821A.raw

L:48 M:34] W: (46) "n" OR "Xaa" used, for SEQ ID#:2